Note Title 2021-09-29

Numeri di macchina normalizzati

$$F(\beta, t, m, M) = \left\{ \beta^p \sum_{i=1}^{t} c_i \beta^{-i} \; : \; 0 \le c_i < \beta \;\; \forall i, \; c_1 \neq 0 \right\}$$

$+0, \; \infty$ $\qquad \beta = 10 \quad t = 3$

$$0.999 = 10^0 \cdot 0.999 = 10^0 \sum \underbrace{9}_{c_1} \cdot 10^{-1} + \underbrace{9}_{c_2} \cdot 10^{-2} + \underbrace{9}_{c_3} \cdot 10^{-3}$$

$m = -2 \quad M = 2$

$$\underbrace{0.1}_{= 10^0 \cdot 0.100} \ominus \underbrace{0.001}_{= 10^{-2} \cdot 0.100} = fl(0.1 - 0.001) = fl(0.099) = \underbrace{10^{-1}}_{exp} \cdot \underbrace{0.990}_{mantissa}$$

$$1 \oslash 3 \otimes 3 = fl(0.33333\cdots) \otimes 3 = (10^0 \cdot \underbrace{0.333}_{= 3 \cdot 10^{-1} + 3 \cdot 10^{-2} + 3 \cdot 10^{-3}}) \otimes 3$$

$$= fl(0.333 \cdot 3) = fl(0.999) = 10^0 \cdot 0.999$$

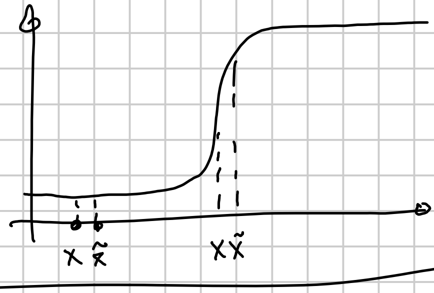
$y = f(x) \qquad f : [a, b] \to \mathbb{R}$ razionale (ottenuta con $+, -, *, /$)

es: $f(x) = x^2 - \frac{3}{x}$

Non posso inserire x come input su un computer, devo approssimarlo con un num. di macchina $\tilde{x}$. Per ogni x t.c. $|x| \in [\omega, \Omega]$ esiste $\tilde{x}$ num. di macchina t.c. $\frac{\tilde{x} - x}{x} = \varepsilon$

$|\varepsilon| < u \qquad \tilde{x} = x(1 + \varepsilon)$

Errore inerente $\qquad \varepsilon_{in} = \dfrac{f(\tilde{x}) - f(x)}{f(x)} \qquad f(x) \neq 0$



$$f(\tilde{x}) = \tilde{x}^2 - \frac{3}{\tilde{x}} \longrightarrow g(\tilde{x}) = \tilde{x} \circledast \tilde{x} \ominus 3 \oslash \tilde{x}$$

$$\varepsilon_{alg} = \frac{g(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})}$$ errore algoritmico

$$\varepsilon_{tot} = \frac{g(\tilde{x}) - f(x)}{f(x)}$$

Teo:

$$\varepsilon_{tot} \doteq \varepsilon_{alg} + \varepsilon_{in}$$

al primo ordine, ignorando termini dell'ordine del prodotto di questi errori

$$\varepsilon_{tot} = \frac{g(\tilde{x}) - f(x)}{f(x)} = \frac{g(\tilde{x}) - f(\tilde{x})}{f(x)} + \frac{f(\tilde{x}) - f(x)}{f(x)} = \varepsilon_{alg} \cdot \boxed{\frac{f(\tilde{x})}{f(x)}} + \varepsilon_{in}$$

Noto che $\varepsilon_{in} = \frac{f(\tilde{x}) - f(x)}{f(x)} = \frac{f(\tilde{x})}{f(x)} - 1 \rightarrow \frac{f(\tilde{x})}{f(x)} = 1 + \varepsilon_{in}$

$$\rightarrow \varepsilon_{tot} = \varepsilon_{alg}(1+\varepsilon_{in}) + \varepsilon_{in} = \varepsilon_{alg} + \cancel{\varepsilon_{alg}\varepsilon_{in}} + \varepsilon_{in} \doteq \varepsilon_{alg} + \varepsilon_{in}$$

prodotto di due errori

Stima al primo ordine dell'errore inerente:

$$\tilde{x} = x(1+\varepsilon) = x + x\varepsilon \qquad |\varepsilon| < u$$

$$f(x+h) = f(x) + f'(x)h + \frac{f''}{2}(\xi)h^2$$

$$f(\tilde{x}) = f(x+\varepsilon x) = f(x) + f'(x)\varepsilon x + \frac{f''(\xi)}{2}\varepsilon^2 x^2$$

$$\varepsilon_{in} = \frac{f(\tilde{x}) - f(x)}{f(x)} = \frac{f'(x)x\varepsilon}{f(x)} + \frac{f''(\xi)}{2f(x)}x^2\varepsilon^2 \doteq \frac{f'(x)x}{f(x)}\varepsilon$$

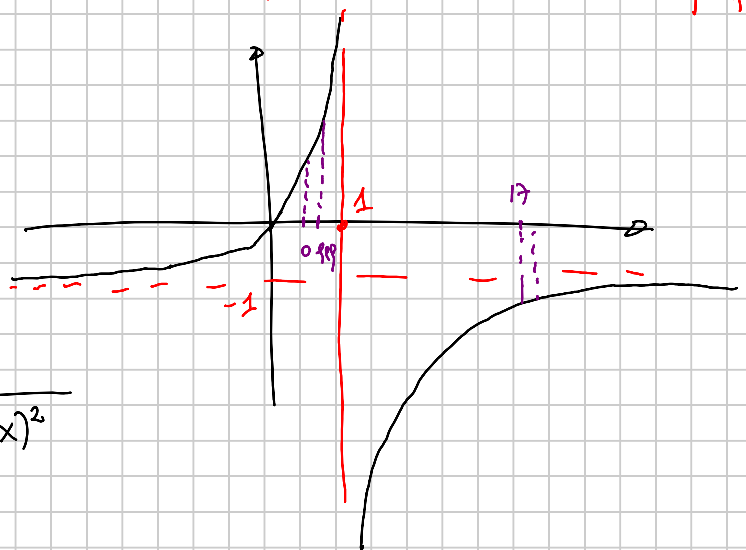
$\varepsilon = \frac{\tilde{x}-x}{x}$ errore <u>relativo</u> sull'argomento x della funzione

$\varepsilon_{in} = \frac{f(\tilde{x})-f(x)}{f(x)}$ errore <u>relativo</u> sul risultato della funzione $\doteq \boxed{\frac{f'(x)\,x}{f(x)}}\,\varepsilon$

numero di condizionamento
(della funzione f nel punto x)

$$K_{f,x} := \left|\frac{f'(x)\,x}{f(x)}\right|$$

<u>ES</u>: $f(x) = \frac{x}{1-x}$.



Numero di condizionamento:

Calcolo prima di tutto

$$f'(x) = \frac{1\cdot(1-x) - x(-1)}{(1-x)^2} = \frac{1}{(1-x)^2}$$

$$K_{f,x} = \left|\frac{\frac{1}{(1-x)^{\cancel{2}}}\,\cancel{x}}{\cancel{\frac{x}{1-x}}}\right| = \frac{1}{|1-x|}$$

se $x \simeq 1$, un piccolo errore (rel.) in x ne causa uno molto grande in $f(x)$

f mal condizionata in $x \simeq 1$

se x lontano da 1, questo non succede.

f ben condizionata in x

$|\varepsilon| \leq u = 10^{-16}$

Stima al primo ordine dell'errore algoritmico

Dato un'espressione $f(\tilde{x}) = \tilde{x}^2 - 3/\tilde{x}$

il computer calcola $g(\tilde{x}) = \underbrace{\tilde{x} \otimes \tilde{x}} \ominus \underbrace{3 \oslash \tilde{x}}$

$\simeq 2.2 * 10^{-16}$ per Matlab

$|\varepsilon_i| \leq u$

$$g(\tilde{x}) = \underbrace{\tilde{x}^2(1+\varepsilon_1)} \ominus \underbrace{\frac{3}{\tilde{x}}(1+\varepsilon_2)} = \left(\tilde{x}^2(1+\varepsilon_1) - \frac{3}{\tilde{x}}(1+\varepsilon_2)\right)(1+\varepsilon_3)$$

$$= \underbrace{\tilde{x}^2 - \frac{3}{\tilde{x}}}_{f} + \tilde{x}^2\varepsilon_1 + \tilde{x}^2\varepsilon_3 - \frac{3}{\tilde{x}}\varepsilon_2 - \frac{3}{\tilde{x}}\varepsilon_3 + \cancel{\tilde{x}^2\varepsilon_1\varepsilon_3} - \cancel{\frac{3}{\tilde{x}}\varepsilon_2\varepsilon_3} \doteq$$

$$\doteq f(\tilde{x}) + \tilde{x}^2\varepsilon_1 - \frac{3}{\tilde{x}}\varepsilon_2 + \left(\tilde{x}^2 - \frac{3}{\tilde{x}}\right)\varepsilon_3$$

$$\frac{g(\tilde{x}) - f(\tilde{x})}{f(\tilde{x})} \doteq \frac{\tilde{x}^2}{\tilde{x}^2 - 3/\tilde{x}}\varepsilon_1 - \frac{3/\tilde{x}}{\tilde{x}^2 - 3/\tilde{x}}\varepsilon_2 + \varepsilon_3$$

$\cancel{|a-b| \geq |a|-|b|}$

$|a+b| \leq |a|+|b|$

$$\left|\frac{g(\tilde{\vec{x}})-f(\tilde{\vec{x}})}{f(\tilde{\vec{x}})}\right| \doteq \left|\frac{\tilde{x}^2}{\tilde{x}^2-3/\tilde{x}}\varepsilon_1 - \frac{3/\tilde{x}}{\tilde{x}^2-3/\tilde{x}}\varepsilon_2+\varepsilon_3\right| \underset{\text{dis. triangolare}}{\leq} \left|\frac{\tilde{x}^2}{\tilde{x}^2-3/\tilde{x}}\varepsilon_1\right| + \left|-\frac{3/\tilde{x}}{\tilde{x}^2-3/\tilde{x}}\varepsilon_2\right| + |\varepsilon_3|$$

$$= \left|\frac{\tilde{x}^2}{\tilde{x}^2-3/\tilde{x}}\right|\cdot|\varepsilon_1| + \left|\frac{3/\tilde{x}}{\tilde{x}^2-3/\tilde{x}}\right|\cdot|\varepsilon_2| + |\varepsilon_3| \leq \underbrace{\left|\frac{\tilde{x}^2}{\tilde{x}^2-3/\tilde{x}}\right| u + \left|\frac{3/\tilde{x}}{\tilde{x}^2-3/\tilde{x}}\right| u + u}_{\text{stima al primo ordine dell'errore algoritmico}}$$

Possiamo applicare la stessa strategia in generale per ogni funzione:

1) scrivo l'algoritmo per calcolare $f(\tilde{\vec{x}})$ con op. di macchina $\oplus, \ominus, \otimes, \oslash$
2) espando $a \,\text{(op)}\, b = (a \text{ op } b)(1+\varepsilon_i)$
3) ignoro termini con due o più ε
4) raccolgo ε_i e prendo valori assoluti

In questo caso, ε_{alg} sarà grande quando $\tilde{x}^2$ e $\frac{3}{\tilde{x}}$ sono vicini fra loro (perché in questo caso $\left|\frac{\tilde{x}^2}{\tilde{x}^2-3/\tilde{x}}\right|$ e $\left|\frac{3/\tilde{x}}{\tilde{x}^2-\frac{3}{\tilde{x}}}\right|$ sono grandi)

$$\tilde{x}^2 \approx 3/\tilde{x}$$
$$\Downarrow$$
$$\tilde{x}^2\cdot\tilde{x} \approx 3 \Rightarrow \tilde{x} \approx \sqrt[3]{3}$$

Terzo tipo di errore: errore analitico

Alcune quantità si possono calcolare solo come limite di processi iterativi, o approssimazioni

es: $\exp(x) = 1+x+\frac{x^2}{2}+\frac{x^3}{3!}+\dots = \sum_{n=0}^{\infty}\frac{x^n}{n!} = \phi(x)$

Dovrò troncare la sommatoria e calcolare

$$f(x) = \sum_{n=0}^{N}\frac{x^n}{n!} = f(x)$$

In generale, per calcolare una funzione $\phi(x)$ la approssimo con una funzione calcolabile con un num. finito di operazioni: $f(x)$

$$\varepsilon_{an} = \frac{\phi(x)-f(x)}{f(x)} \qquad \varepsilon_{in} = \frac{f(\tilde{x})-f(x)}{f(x)} \qquad \varepsilon_{alg} = \frac{g(\tilde{x})-f(\tilde{x})}{f(\tilde{x})}$$

$$\frac{\phi(x)-g(\tilde{x})}{\phi(x)} \doteq \varepsilon_{an}+\varepsilon_{in}+\varepsilon_{alg}$$

Spesso la comp. dominante della somma (molto più grande degli altri) sarà l'errore analitico.

Condizionamento delle quattro operazioni:

<u>prodotto</u>:

$$x, y \to xy$$

Studiamo perturbazioni $\tilde{x} = x(1+\varepsilon_x)$ e $\tilde{y} = y(1+\varepsilon_y)$ contemporaneamente

$$\tilde{x}\tilde{y} = x(1+\varepsilon_x)y(1+\varepsilon_y) = xy + xy\varepsilon_x + xy\varepsilon_y + xy\varepsilon_x\varepsilon_y$$

$$\doteq xy + xy\varepsilon_x + xy\varepsilon_y$$

$$|\varepsilon_x, \varepsilon_y| < u$$

$$\varepsilon_{in} = \frac{\tilde{x}\tilde{y} - xy}{xy} \doteq \frac{xy\varepsilon_x + xy\varepsilon_y}{xy} = \varepsilon_x + \varepsilon_y$$

$$|\varepsilon_{in}| = |\varepsilon_x + \varepsilon_y| \leq |\varepsilon_x| + |\varepsilon_y| \leq 2u$$

<u>somma, sottrazione</u>

Ci basta studiare la somma, visto che $x - y = x + (-y)$

$$\tilde{x} = x(1+\varepsilon_x) \qquad \tilde{y} = y(1+\varepsilon_y)$$

$$\tilde{x} + \tilde{y} = x(1+\varepsilon_x) + y(1+\varepsilon_y) = x + y + x\varepsilon_x + y\varepsilon_y$$

$$\varepsilon_{in} = \frac{\tilde{x}+\tilde{y} - (x+y)}{x+y} = \frac{x\varepsilon_x + y\varepsilon_y}{x+y}$$

$$|\varepsilon_{in}| = \left| \frac{x}{x+y}\varepsilon_x + \frac{y}{x+y}\varepsilon_y \right| \leq \underbrace{\left|\frac{x}{x+y}\right|}u + \underbrace{\left|\frac{y}{x+y}\right|}u$$

può essere arbitrariamente grande se x e y sono molto vicini ed essere uguali e opposti $x = 5$, $y = -4.999999$

$$\left|\frac{x}{x+y}\right| = \frac{5}{0.000001} = 5 \cdot 10^6$$

<u>divisione</u>: $\tilde{x} = x(1+\varepsilon_x)$ $\qquad \tilde{y} = y(1+\varepsilon_y)$

$$\frac{\tilde{x}}{\tilde{y}} = \frac{x(1+\varepsilon_x)}{y(1+\varepsilon_y)} = \frac{x}{y}(1+\varepsilon_x)(1 - \varepsilon_y + \varepsilon_y^2 - \varepsilon_y^3 + \ldots) =$$

$$(1+\varepsilon_y)^{-1} = 1 - \varepsilon_y + \varepsilon_y^2 - \varepsilon_y^3 + \varepsilon_y^4 - \ldots$$

serie geom. di ragione $-\varepsilon_y$

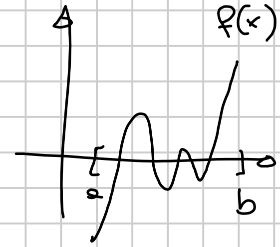
$\varepsilon_{in} \doteq \frac{x}{y}\cdot 1 + \frac{x}{y}\varepsilon_x - \frac{x}{y}\varepsilon_y$ se $|\varepsilon_x|, |\varepsilon_y| < u$

$$\frac{\tilde{x}/\tilde{y} - x/y}{x/y} \doteq \frac{\frac{x}{y}\varepsilon_x - \frac{x}{y}\varepsilon_y}{x/y} = \varepsilon_x - \varepsilon_y \qquad |\varepsilon_{in}| = |\varepsilon_x - \varepsilon_y| \le |\varepsilon_x| + |\varepsilon_y| < u + u = 2u$$

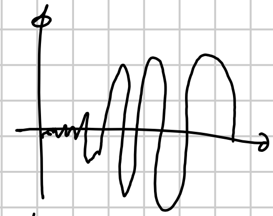
Problema: soluzione numerica di equazioni non lineari

Data $f: [a,b] \to \mathbb{R}$

voglio trovare un punto x_* tale che $f(x_*) = 0$



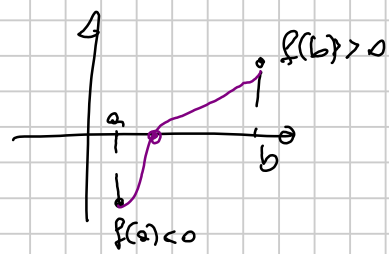
Come faccio ad assicurarmi che esista almeno una soluzione nell'intervallo $[a,b]$?



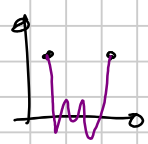
Una condizione semplice da verificare:

Se $f \in C^0([a,b])$ e $f(a)$, $f(b)$ hanno segni opposti allora $\exists\, x_* \in [a,b]$ t.c. $f(x_*) = 0$

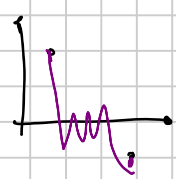
("teorema degli zeri")



non è un se e solo se



non assicura zero unico



Un modo di assicurarsi che lo zero sia unico è dimostrando per esempio che la f è strettamente crescente o decrescente

(se f derivabile, basta mostrare $f'(x) > 0 \ \forall x \in [a,b]$
oppure $f'(x) < 0 \ \forall x \in [a,b]$)

Def: $[a,b]$ si dice intervallo di separazione per f se $f(a)f(b) < 0$.
(cioè, $f(a)$ e $f(b)$ hanno segno opposto).